

DOCUMENT RESUME

ED 268 248

CE 043 933

AUTHOR Carlisle, Ysanne
 TITLE Descriptive Statistics. Supervising: Technical Aspects of Supervision. The Choice Series #43. A Self Learning Opportunity.
 INSTITUTION Northern Regional Management Centre, Washington (England).
 SPONS AGENCY Manpower Services Commission, London (England).
 REPORT NO ISBN-0-948023-46-5
 PUB DATE 85
 NOTE 56p.; For related documents, see ED 266 310-312 and CE 043 921-948.
 AVAILABLE FROM Northern Regional Management Centre, Derwent House, Washington Town Centre, Washington, Tyne & Wear NE38 7ST, England (\$15.00 plus postage; 15% discount on 50, 25% on 100; includes audiotape).
 PUB TYPE Guides - Classroom Use - Materials (For Learner) (051)
 EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
 DESCRIPTORS Adult Education; Behavioral Objectives; Learning Activities; Management Development; Self Evaluation (Individuals); *Statistical Distributions; *Statistics; Student Evaluation; *Supervision; Supervisor Qualifications; *Supervisory Methods; *Supervisory Training
 IDENTIFIERS *Data Organization; Great Britain

ABSTRACT

This learning unit on descriptive statistics is one in the Choice Series, a self-learning development program for supervisors. Purpose stated for the approximately eight-hour-long unit is to enable the supervisor to understand the nature and uses of statistics at work, understand information which is obtained from figures at work, use statistics to present and describe numerical data, and form judgments by "reading" various kinds of data presentations and statistics. An introduction provides an overview of the unit and lists unit objectives. Parts A through D contain informative material, activities and answer keys, self-checks and answer keys, and a summary. Topics include uses and nature of statistics, "on the average," grouped data, and the spread values. Part E contains performance checks--an end check to be completed alone and a tutor check and work-based project. Part F consists of a brief unit review, extensions (print sources of additional information, such as units or sections in books), and a list of references. (YLB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

An
Open Tech
Project

SUPERVISING

TECHNICAL ASPECTS OF SUPERVISION

43: Descriptive Statistics

ED268248

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
- ☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

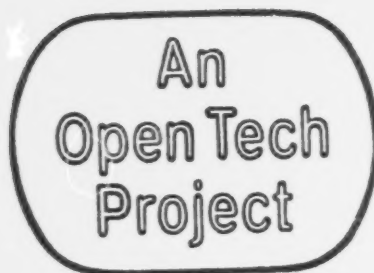
"PERMISSION TO REPRODUCE THIS
MATERIAL IN MICROFICHE ONLY
HAS BEEN GRANTED BY

H. Harlow

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

43: DESCRIPTIVE STATISTICS

Ysanne Carlisle



The CHOICE programme is an Open Tech sponsored project managed by the Northern Regional Management Centre in conjunction with the Northern Consortium of Colleges and the National Examinations Board for Supervisory Studies.

© Crown Copyright 1985 Published by permission of the Controller
of Her Majesty's Stationery Office.

ISBN 0 948023 46 5

First published in the U K - 1985

Northern Regional Management Centre
Derwent House, Washington Town Centre
Washington, Tyne & Wear NE38 7ST

This Work was produced under an Open Tech contract with the Manpower
Services Commission. The views expressed are those of the authors and
do not necessarily reflect those of the MSC or any Government
Department.

All rights reserved. No part of this book may be reproduced or
transmitted in any form or by any means, electronic or mechanical,
including photocopying, recording or by any information storage and
retrieval system, without permission in writing from the publisher.

Ysanne Carlisle
Descriptive Statistics (Number 43 in the CHOICE SERIES of
learning units)

Series Editor: Harry V Pardue
Northern Regional Management Centre.

CONTENTS

Page

PART THE USES AND NATURE OF STATISTICS

1	Introduction	2
2	Using statistics in business	2
3	The nature of statistics	3
4	Tables	5
5	Summary	10

PART B "ON THE AVERAGE"

1	Introduction	11
2	Three 'statistics'	11
3	The mean	12
4	Frequency distribution tables	16
5	Summary	22

PART C GROUPED DATA

1	Introduction	23
2	Grouping data	23
3	Calculating the mean for grouped data	26
4	Summary	28

PART D THE SPREAD VALUES

1	Introduction	29
2	The range	29
3	Modifying the range	31
4	The mean deviation	34
5	Summary	39

PART E PERFORMANCE CHECKS

1	End check	40
2	Tutor check	44
3	Work-based project	45

PART F UNIT REVIEW

1 Return to objectives	46
2 Extensions	47
3 References	48

WORKING GUIDE TO THIS UNIT IN YOUR DEVELOPMENT PROGRAMME

This Unit

is now your personal property - it's not for handing back to your training officer or to a tutor. It is your working book.

The Contents

will increase in value to you each time you write on one of its pages. This is a working text for your personal notes - without notes its value is limited.

Qualifications

forget about certificates and diplomas - you don't need them for this programme. What you do need is the wish:-

- * to do your job better
- * for personal development

Your study time

is really up to you - we guess you will need about 8 hours in total. But it's your time - spend as long as you like.

How am I doing?

you can judge for yourself by checking out how well you cope with the tasks we call Activities and Self Checks which occur throughout the unit. You can also get the Tutor Check and/or the Work Based Project to your tutor for an unbiased view of your progress.

Work Pace

it's really up to you and depends on your interest. However, if you do well on the Activities and Self Checks, you could afford to speed up a little. Of course, if you don't do as well as you expect then you could slow down a little or flick back into the unit for a quick review of the more difficult ideas.

When in doubt

- ask or seek help from:

- * your family and friends
- * colleagues at work
- * your company training staff
- * other open learners
- * your tutor, by phone, by a visit to your support centre or by dropping your tutor a line through the post.

Don't worry

if what you learn doesn't seem to be of use to you now, remember that times change - it may be useful one day. Anyway, just knowing more improves personal confidence.

So now make a start:-

43: DESCRIPTIVE STATISTICS by Ysanne Carlisle

ABOUT THIS UNIT

Numerical information or data is produced and used by almost every area in a business organisation. Statistics is a subject which can help us to examine such information in a meaningful way. There is a skill in dealing with numbers, and managers at all levels, from the supervisor to the managing director, need to acquire some knowledge in this area if they are to make any real sense of these facts and figures. Figures should not be used to support just any argument. In the popular press, statistics are often misused to support views which a statistician would know they don't justify. Misuse or misinterpretation of numerical information in a business can lead to poor decisions and damaging consequences. In this unit we will try to develop an awareness of statistics as a tool, what statistical data can and cannot be used for and how you as a supervisor can draw on an understanding of statistics to help you in your day-to-day work.

IN THIS UNIT WE WILL:

- * develop an awareness of the scope and nature of statistics
- * explore some of the different ways of using statistics to present and describe numerical information
- * develop an ability to extract useful measures from sets of data
- * explore the ways of looking at statistics produced at work

AUDIO TAPE Side 1

Before you continue through the unit, you should listen to side 1 of the audio tape that accompanies the unit. You will find it useful if you refer to the unit objectives shown below whilst you listen to the tape. You may wish to make the objectives clearer to yourself by making a note or two on the page as you listen.

OBJECTIVES

After you've completed this unit you will be BETTER ABLE to:

- * understand the nature and uses of statistics at work
- * understand information which is obtained from figures at work
- * use statistics to present and describe numerical data
- * form judgements by 'reading' various kinds of data presentations and statistics

PART A THE USES AND NATURE OF STATISTICS

1 Introduction

As a supervisor you have to handle a lot of facts and figures about your job and the performance of your workteam.

These facts and figures may be referred to as statistics - but the word 'statistics' has a number of meanings.

It may mean simply a number, any number, or a measurement. In this unit, we consider statistics as a collection of techniques which we can use to make sense of numbers - calculating an average is an example of such a technique. We can make use of these techniques to help us describe and interpret figures. They are as useful to us for this purpose as a hammer is for knocking in nails.

Should you be required to collect information of a numerical kind, a basic knowledge of statistics will help in your task of collecting and presenting it effectively. If someone else gives you some figures to look at, statistics will help you to understand what's going on and what needs doing.

2 Using Statistics in Business

ACTIVITY 1
Time guide
2 mins

Write down TWO examples of numerical information you look at as part of your job. Write your responses in the space provided.

* Number of items in stock

* _____

* _____

Typical responses might include:

- * number of pieces of spoiled work in your section
- * output figures
- * number of hours lost through machine breakdown or downtime
- * number of hours lost through absenteeism in your section

These are just a few of the possibilities. However statistics can be used to examine all this information and much more throughout the organisation. Peter Clark writes about some of these in his book "Using statistics in business - 1".

REFERENCE 1
page 48

We can use statistics:

- * to analyse production costs, output levels and defect rates
- * in work and methods study
- * in quality control and inspection
- * in market research
- * in forecasting future sales
- * in comparing the performances of alternative products
- * to examine wage rates and changes
- * in looking at absenteeism
- * in investment decisions
- * in manpower planning
- * in cost control

This list of possible uses, like the list of numerical information we looked at earlier, is endless. However, from this short exercise we can see how wide the use of statistics is. Every department, from Production to Personnel, Accounts to Marketing, produces numerical information to which we can apply statistics.

By now you've probably got some idea about how widely used statistics is.

Let's consider the nature of our tool and what it can do for us.

3 The Nature of Statistics

Statistics can be thought of as a science of numbers which helps us to present, describe and interpret numerical data.

SELF CHECK 1
Time guide
1 min

Look at each of the following areas of information. Put a tick beside the ones which you think can be described numerically.

- * Job satisfaction experienced by members of your workteam
- * The output of your section
- * The mood of the members of your section
- * The defect rates on your section
- * Attitudes towards a job improvement programme
- * How the bonus payments on your section compare with those on others.

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. Output and defect rates are easily measured, we could put a figure on both. The same can be said of bonus payments - it would be easy to compare total payments made to various sections. Bonus payments can be expressed numerically in pounds and pence.

DATA WHICH CAN BE EXPRESSED NUMERICALLY IS
CALLED QUANTITATIVE DATA

It's difficult to see how we can put a numerical value on the other areas of information. Job satisfaction, or lack of it, is purely personal experience relating to an individual quality of working life. We can say for example that an individual is happy or sad but we can't readily express either mood in terms of numbers. We can say the same things about attitudes. Someone may have a positive or negative attitude towards something but we can't put a value on it.

DATA WHICH CAN'T BE EXPRESSED NUMERICALLY IS
CALLED QUALITATIVE DATA

So this self check illustrates what statistics is, and is not, concerned with.

STATISTICS IS CONCERNED WITH QUANTITATIVE OR
NUMERICAL DATA

STATISTICS IS NOT CONCERNED WITH QUALITATIVE
DATA

4 Tables

We may find that data when first collected is not arranged in a way which suits our needs. For example, look at this next case incident:

ACTIVITY 2
Time guide
5 mins

Let's assume that a new product has been introduced in a company where the operators are paid a bonus for output above a standard level. The standard hasn't been settled yet. In one of the sections involved the operators are required to turn out simple brackets to be used in a later stage of production. In order to settle a standard, the operators were given training, then a work study engineer came to observe eight different operators during five hourly trials over a period of two days. He found they had completed the following number of brackets on the trials.

TABLE OF NUMBERS OF BRACKETS PRODUCED

	1st Hour	2nd Hour	3rd Hour	4th Hour	5th Hour
OPERATOR A	2	3	4	4	4
OPERATOR B	1	2	3	3	4
OPERATOR C	3	4	4	4	5
OPERATOR D	3	4	3	4	4
OPERATOR E	2	3	3	5	6
OPERATOR F	2	2	3	2	3
OPERATOR G	3	4	3	3	4
OPERATOR H	2	3	3	2	3

If we were the work study officer we'd want to know how many brackets an operator can reasonably be expected to make in an hour - this would be our standard level.

So it would be better for our purposes if we had a table showing the number of times that each actual output level was achieved. Let's do this using the information given here. I've put in the first two entries:

How frequently was ONE bracket produced	1
How frequently were TWO brackets produced	8
How frequently were THREE brackets produced	
How frequently were FOUR brackets produced	
How frequently were FIVE brackets produced	
How frequently were SIX brackets produced	

The number of brackets produced in each of the hours varied from 1 to 6. It's time consuming and unnecessary to write "How frequently was A NUMBER of items produced" on each line, so instead we could simply give each of the two columns a title and list the numbers. Had we done this our table would have looked like this:

Number of brackets produced on the trials	The frequency of the numbers produced
1	1
2	8
3	16
4	12
5	2
6	1
TOTAL	40

You can check your work against the right-hand column of this table. Note that the total of the right hand column adds up to 40, that's because the work study engineer conducted 40 trials.

So, in this activity we just used our first statistical method, TABULATION, drawing up a table which arranged our data in a more convenient way. It is now easy to see, for example, that more operators produced 3 brackets in an hour than any other number.

Statisticians would use the letter 'x' for the number of items produced in an hour, it's useful shorthand for describing the actual values which are of interest to us. They would also use the letter 'f' for the frequency with which each of those values occurred. Columns in tables would therefore be headed x and f.

ACTIVITY 3
Time guide
mins

Tabulate the figures from Activity 2, using the conventional symbols x and f to head the columns. I've completed the first two lines.

x	f
1	1
2	8
* —	* —
* —	* —
* —	* —
* —	* —
—	—

Our completed table will look like this:

x	f
1	1
2	8
3	16
4	12
5	2
6	1

You've just completed a FREQUENCY DISTRIBUTION table, which is a table showing the values in which we are interested in the left hand column and the frequencies with which they occur in the right hand column. As Peter Betts points out, frequency distributions are a way of presenting a mass of data more simply and more usefully. We'll see that this is true in later parts of the unit.

SELF CHECK 2 1. Complete the sentences below with a suitable word or words:

Time guide
5 mins

- a) Statistics is the science of numbers which helps us to present, describe and interpret _____ data.
- b) Arranging data in a helpful way to suit our needs involves the technique of _____.
- c) A table which shows the values in which we are interested in the left hand column and how frequently they occur in the right hand column is called a _____ table.
2. 10 encyclopaedia salesmen managed to sell the following number of encyclopaedia sets in each of 5 weeks:

Number of sets sold in each of 5 weeks					
Salesman A	1	2	2	3	4
B	2	2	3	2	3
C	3	3	4	4	3
D	2	3	4	5	6
E	2	2	3	3	4
F	2	3	4	5	6
G	2	2	3	2	2
H	1	2	2	3	5
I	1	2	3	3	3
J	3	4	4	5	6

Complete a frequency distribution table for these figures.

The answers to this self check are here in the following words, so you'll need to check out your answers carefully. 1 a) Quantitative. The word numerical would also have been appropriate here. This highlights the distinction between quantitative data which is a concern of statistics and qualitative data which is not ; b) Tabulation or drawing up a table. c) Frequency Distribution.

2.

Number of encyclopaedia sets sold in a week	frequency
x	f
1	3
2	16
3	16
4	8
5	4
6	3

50

Note that in this example the total of the frequency column is 50. This is because the original table recorded 5 figures for each of 10 salesmen. There were 50 recorded results.

We will be using frequency distribution tables again later on but before we do, we should note that we can draw up frequency distribution tables dealing with a wide range of subjects.

Here are some examples:

Number of days absent from work in a year	Number of people
x	f
1	20
2	40
3	65

Amount of bonus earned £'s	Number of people
x	f
1	1
2	10
3	20

Number of cigarettes smoked each day	Number of people	Number of visits made to a restaurant in a month	Number of people
10	0	1	10
20	2	2	12
30	5	3	11

5 Summary

Before moving on, let's review our work so far.

- * STATISTICS are commonly and widely used in business
- * STATISTICS IS A SCIENCE OF NUMBERS which:
 - is concerned with QUANTITATIVE DATA
 - helps us to PRESENT, DESCRIBE and INTERPRET NUMERICAL DATA
- * DATA COLLECTED can be arranged to suit our needs:
 - TABULATION is the technique we use to arrange data in a suitable way
 - FREQUENCY DISTRIBUTION TABLES are tables which list the values in which we are interested and the frequencies with which they occur.
- * Frequency Distribution Tables are useful for 'condensing' a lot of data for ease of use.

PART B "ON THE AVERAGE"

1 Introduction

In this part of the unit we'll look at what we mean by "on the average". In everyday life we use the word average to describe lots of different things - our average weekly pay, our average shopping bill, our average mortgage. We read in the papers about the average number of children, the average wage, the average house - the list is endless. But what do we use the average for?

Well it's really some single figure which we hope represents a group of figures.

But how do we get at this one figure which we use to generally describe the group of figures? The answer is not easy, because there are several ways of doing this. Let's see what they are.

2 Three 'statistics'

ACTIVITY 4
Time guide
2 mins

Look at this frequency distribution table - you'll probably recall it's the one we used earlier.

Number of brackets produced per hour	frequency
x	f
1	1
2	8
3	16
4	12
5	2
6	1

40 - our check figure

What single figure for "brackets produced in an hour" best describes those figures in our 'x' column.

* _____ brackets per hour

Of course there are a number of answers here and no doubt we could justify them all, but perhaps the three most likely figures are:

- * $2\frac{3}{4}$ - half the values lie below this figure and, of course, half lie above this figure.
- * 3 - this is the number most often achieved - this number was made 16 times
- * $3\frac{1}{4}$ - this is just about the average number produced during the whole trial

But how do we really decide? Let's deal with this question by identifying the three possible statistics which can be used to represent all the figures considered. In this example:

$2\frac{3}{4}$ is the median

3 is the mode

$3\frac{1}{4}$ is the mean

EXTENSION 1
page 47

In this unit we are not going to explore the MODE or the MEDIAN. If you wish to know more about the MODE then why not take up this extension.

The mean is the most commonly used statistic for describing a large number of figures. It is an important statistic because it helps us to calculate further and makes planning for the future possible. The mean also lends itself to a whole range of other statistics all very useful to us at work.

So let's move on to see just how useful the mean is.

3 The Mean

Now let's take a look at the second likely response to Activity 4. You will recall this was $3\frac{1}{4}$ and we know this is the MEAN.

We can set about calculating the mean by adding up the total number of brackets and dividing this total by the total number of trials. We then get the mean number of brackets produced in an hour.

ACTIVITY 5
Time guide
2 mins

Let's do this just as we have described it here. You can fill in the figures and do the calculations. Go back to our "table of number of brackets produced" in Activity 2 to get the figures you need:

Total number of brackets produced = * _____

Total number of trials = * _____

So our mean is * _____ divided by * _____ = * _____

Here are the figures to check out our calculation: total number of brackets produced is 129, total number of trials is 40, so our mean is 129 divided by 40 which is just about $3\frac{1}{4}$.

We can calculate the mean for many kinds of things. Some examples might include

- The mean number of man hours lost in a week through machine breakdowns
- The mean number of days lost by each man in a year through sickness
- The mean commission earned by the members of a sales team

ACTIVITY 6
Time guide
1 min

Write down in the space provided below two examples of means which could be calculated from data at your workplace.

*

*

In addition to our three earlier examples we could note these:

- * The mean number of pieces of work spoiled on the section in a week
- * The mean number of enquiries dealt with in a local government office in a week
- * The mean number of calls to customers made by a sales representative in a day
- * The mean number of items of a particular type in stock

The list of possible means we could calculate is endless.

You will have noticed we have used the word 'mean' all the time - you might be more used to the word 'average'. It really doesn't matter too much as long as we understand that we are referring to the same statistic.

SELF CHECK 3 In each of the following questions, calculate the means.
Time guide
3 mins

- 1 60 items were produced in 5 hours. Mean = * _____
 - 2 200 salesmen made 700 sales in a day. Mean = * _____
 - 3 £160 was earned in bonus payments by 50 people. Mean = * _____
-

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. In each of the cases, we calculate our average by dividing the total of the values in which we are interested by the total number of values. 1) 60 items divided by 5 hours gives a mean of 12 items per hour; 2) 700 sales divided by 200 salesmen gives a mean of $3\frac{1}{2}$ sales per salesman a day. 3) £160 divided by 50 people gives a mean bonus of £3.20 per person.

EXTENSION 2 In calculating the mean there are times when you will need to
Page 47 approximate. In other words, write down a sensible figure rather than the actual figure. If we calculate an average bonus to be £4.1242p it would be sensible to give the answer as £4.12p, rather than £4.1242p. This extension gives you an opportunity to explore the process of approximation further. You'll find we approximate in other parts of this unit too. So long as our approximates are suitable for our purposes, there is no harm in it.

3.1 A word of caution

There are some circumstances where the mean is not an accurate description of a set of figures. We'll need to know about this so look at this next activity:

ACTIVITY 7
Time guide
5 mins

The number of invoices dealt with by a typist in each of two weeks was as follows:

	Monday	Tuesday	Wednesday	Thursday	Friday
Week 1	160	175	200	180	170
Week 2	170	180	160	175	80

Calculate the mean weekly output in each of the two weeks by filling in the blanks below

Week 1

Total number of invoices handled was * _____

Total number of days involved was * _____

Mean weekly output was * _____

Week 2

Total number of invoices handled was * _____

Total number of days involved was * _____

Mean weekly output was * _____

During week 1 the typist handled 885 invoices in 5 days so the mean is 885 divided by 5 or 177 invoices per day.

During week 2 the same typist handled 765 invoices in 5 days so the mean is 765 divided by 5 or 153 invoices per day.

How did this difference in the means arise?

- we note that $177 - 153$ is 24 which is a large difference in the means.

On looking at the figures again we see that in week 1 on Wednesday 200 invoices were handled and that in week 2 on Friday only 80 invoices were handled. Otherwise the outputs in the other days in each week comes to the same total and of course the same mean.

The lesson we draw is that:

THE MEAN CAN BE UNREASONABLY AFFECTED BY EXTREME VALUES

Before including an extreme value in our calculations we should check out how such values may have arisen. For example perhaps the typist had some help on Wednesday in week 1 and may have been absent for a half day on Friday of week 2.

If this was the case we would be quite justified in ignoring these extremes and calculate the averages on the remaining figures. In this case the average in week 1 would be 685 divided by 4 = just over 171 and the average for week 2 would also be just over 171 - the means would then be the same.

In Part D section 3 we can see how we can handle this difficulty by attending to these extreme values before calculating the mean.

You may recall an earlier remark where we said the mean helps us to move on to produce further useful statistics. Well let's see how we may do this by using a special table which makes it easier to handle a lot of figures.

4 Frequency Distribution Tables

We actually produced a frequency distribution table earlier in our unit - though we didn't use it for any other purpose than to 'condense' the data. This is of course useful in itself but frequency distribution tables have other uses, for example, we can calculate the mean much quicker and with less chance of error.

Let's return to our familiar example of the work study engineer's tests in Part A.

Number of brackets produced per hour	frequency
x	f
1	1
2	8
3	16
4	12
5	2
6	1
40 - our check figure	

We said in section 4 that the total of the frequency column was 40, because the work study engineer had carried out 40 observations. This total served as a useful check to make sure that we had included all the results in our table. When we are dealing with large numbers of things, for example

- 40 individual observations of one hour each
- The number of days absent for each of 200 people
- The bonus payments received by each of 150 operators

we can check that we have not accidentally missed out one of our results. This is an important point when we are working out the mean.

SELF CHECK 4 The attendance records of ten people working in an office were examined
Time guide to discover the mean number of working days lost in one month. The
3 mins number of days lost were as follows

Person A was absent for 2 days
Person B was absent for 1 days
Person C was absent for 3 days
Person D was absent for 2 days
Person E was absent for 1 days
Person F was absent for 2 days
Person G was absent for 5 days
Person H was absent for 3 days
Person I was absent for 0 days
Person J was absent for 1 days

Calculate the mean number of days lost per person

* _____ divided by * _____ = * _____

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. The mean number of days lost in the month is calculated by dividing the total number of days lost in the month by the total number of people in the office. So the mean number of days lost in the month is: Total number of days lost was 20 days; The total number of people was 10; mean number of days lost is therefore 20 divided by 10 = 2 days.

Let's suppose that while we were adding up the total number of days lost we missed out the days lost by person G in our addition. Then our mean number of days lost would have been different. Total number of days lost would be 15; Total number of people involved was 10. The mean number of days lost would therefore have been calculated as 15 divided by 10 which is $1\frac{1}{2}$ days.

The difference between 2 and $1\frac{1}{2}$ is important and was caused by a simple error. It can be easy to miss out a figure when we are adding up long lists.

Our mistake is not hard to spot in this example as we only have 10 figures to deal with. But sometimes we could have 30 or 40 or even hundreds of figures - mistakes would be hard to spot in those cases.

DRAWING UP FREQUENCY DISTRIBUTION TABLES
AND USING THEM TO CALCULATE THE MEAN
HELPS AVOID ERROR

SELF CHECK 5
Time guide
2 mins

Complete the sentences with a suitable word or words.

- 1 The _____ is the value which occurs most frequently in a group of figures.
- 2 Half of our values lie above the _____ and, of course, half lie below.
- 3 The _____ is the total of all values divided by the number of values.
- 4 A _____ table helps us to avoid making simple errors.

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. 1) The mode occurs most frequently. 2) half of the figures lie above the median and, of course, half lie below. 3) this calculation gives the mean. 4) the frequency distribution table helps us to handle figures without error.

Now that we have seen how frequency distribution tables can help us avoid the mistake of missing a value out, we will use the table to work out the mean. Here is our earlier frequency distribution table:

Number of brackets produced per hour	frequency
x	f
1	1
2	8
3	16
4	12
5	2
6	1
	40

So how do we start? Well we can see for instance that there were 12 occasions when 4 brackets were produced in an hour. So how many brackets is this? Well the answer is 48 or 4×12 . We can see that this is

"number of brackets produced per hour" $\times$ "frequency"

We use a shorthand way of describing this: it is fx . Therefore we can add a third column to our frequency distribution table, headed ' fx '. Let's do this and put our 48 in as we have done this calculation already.

ACTIVITY 8
Time guide
2 mins

Complete the table by multiplying each value for ' x ' with its corresponding ' f ' value. Then add all the fx values and enter the total at the bottom:

Number of brackets produced per hour	frequency	
x	f	fx
1	1	* ____
2	8	* ____
3	16	* ____
4	12	* 48
5	2	* ____
6	1	* ____
Totals	40	* ____

Our completed table will look like this.

Number of brackets produced per hour	frequency	
x	f	fx
1	1	1
2	8	16
3	16	48
4	12	48
5	2	10
6	1	6
	40	129

We can see that the total of our fx column, 129, is the total number of brackets produced altogether. If we look back to Activity 2, we'll see it's the same. The total of our f column, 40, is the total number of trials - so we can now see that our mean is 129 divided by 40 - very nearly 3 $\frac{1}{4}$. When using our frequency distribution table:

Mean = total of fx divided by total of f

Or sum of fx divided by sum of f

If you ever see the Greek symbol $\sum$ (sigma) in statistics it means "sum of" so if we were to use this symbol $\sum$ then:

Mean = $\sum fx$ divided by $\sum f$

Once we get used to the frequency distribution table we'll soon find it easier to calculate our mean from it rather than go back to the original mass of figures.

SELF CHECK 6
Time guide
7 mins

Let's look at another example of a frequency distribution table of the number of accidents in a week over a large number of weeks:

x	f	fx
1	8	8
2	11	22
3	6	18
4	17	68
5	10	50
6	12	72
7	6	42
TOTALS	70	280

Use this table to complete the questions in this self check.

Respond to each of the multiple choice questions by putting a tick against the ONE answer which is correct.

- (1) The total number of weeks recorded in the frequency distribution table is
 - (a) 7
 - (b) 70
 - (c) 280
- (2) The total number of accidents recorded in the frequency distribution table is
 - (a) 7
 - (b) 70
 - (c) 280
- (3) The mean number of accidents in a week is
 - (a) 4
 - (b) 7
 - (c) 10

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. 1) The total of the frequency column tells us the total number of weeks in the table. In this case the number of weeks recorded was 70; 2) The total number of accidents is 280, the total of the fx column; 3) We decided that the total number of accidents is 280 and that the total number of weeks recorded was 70. The mean number of accidents is 4 and is found by dividing 280 by 70.

EXTENSIONS
3 and 4
page 47

We've looked at how to present data in the form of a frequency distribution table. Extensions 3 and 4 invite you to look at some other methods of presenting data not covered in this unit.

5 Summary

- * We discovered 3 statistics, each of which might be used to generally represent a group of values:

Mean - the total of all values divided by the number of values

Mode - the value which occurs most often

Median - the value at that point which splits our number of values in half.

- * We also decided that the mean was to be our main statistic.
- * When calculating our mean we can use sensible approximates eg 6 13/16 becomes 6½; 3.28 becomes 3.3.
- * Our note of caution about the mean was that it could be affected unreasonably by extreme values and that we can with caution omit the extreme value before calculating the mean.
- * Frequency distribution tables were useful for:
 - * condensing data
 - * calculating the mean
 - * avoiding errors when calculating the mean
- * Using our frequency distribution table:
 - * the mean = Sum of fx divided by Sum of f
 - or, remembering our symbol sigma Σ
 - the mean = Σfx divided by Σf

PART C GROUPED DATA

1 Introduction

So far, we've looked at data which takes whole number values - 4, 7, 3, 18 are whole numbers but $4\frac{1}{4}$, $7\frac{1}{2}$, 8.1, 9.7 are not. However, much of the data we deal with will not take whole number values. People don't always earn an exact number of pounds. A factory will not always produce an exact number of tons of steel girders. It's quite possible to collect 40 values for machine downtime or idle time in a particular plant and for every one of these values to be different. Even if we are dealing with whole number values, we may have figures which vary so widely, say 1 to 500, that it would be time consuming to put each individual value in a table.

2 Grouping Data

As we mentioned in the introduction it is quite possible for every value in a distribution to be different. Even if our values take whole numbers, we may have so many that it would be very difficult to use a table listing each one. We can get over this problem by arranging our values in groups. If we are going to do this, we must be clear about how we define our groups. Before we look at an example, let's think about the kinds of data we might need to arrange in groups.

ACTIVITY 9 Time guide 2 mins

Write down TWO examples of numerical information which might best be arranged in groups:

* The earnings of a large number of people

* _____

* _____

Typical responses might have included:

- * The number of man hours lost per week through machine or equipment breakdowns
 - * Output of a section measured by volume or weight
 - * Sales in pounds
 - * Dates of birth of staff
- and other items.

In the first aid centre of a large office block the nursing staff keeps a record of the total number of visits made to the centre by the employees.

Because this number varies between 1 and 50 we will find it convenient to group this data. We could use 8 or 10 groups - in our case 10 groups is more convenient.

number of visits	frequency of this number of visits
x	f
1 - 5	3
5 - 10	6
10 - 15	12
15 - 20	14
20 - 25	20
25 - 30	17
30 - 35	16
35 - 40	10
40 - 45	4
45 - 50	2
	104

Now if one of our 'x' values was 5, where would we record it? There are two possible places:

in the 0 - 5 group

or in the 5 - 10 group

It doesn't really matter as long as we are consistent when handling data which falls on 5, 10, 15 and so on.

So we will adopt the following practice:

0 - 5 group means - more than 0; up to 5

5 - 10 group means - more than 5; up to 10 and so on.

Therefore our 'x' item of 5 would be recorded in the 0 - 5 group.

So we can now see how we can group data and we know how to handle values which fall on our group extremes.

ACTIVITY 10
Time guide
5 mins

Below we've a set of 45 numbers. Let's condense it by making out a frequency distribution table.

1.5	5.1	2.9	9.8	14.3
19.9	1.8	25.0	28.2	23.2
32.0	40.0	16.8	4.1	30.2
5.9	10.2	3.6	14.9	9.1
34.1	29.8	13.6	12.4	15.2
8.1	5.6	24.4	4.9	13.8
30.0	14.0	18.6	11.3	30.0
13.2	15.1	7.3	20.4	14.7
24.3	19.4	15.4	17.9	22.4

We recall that frequency distribution tables condense data and make it easier to handle. If we could find a way of grouping our 45 figures we could draw up a frequency distribution table which wouldn't be too big. If we look at our figures, the highest is 40.0. The lowest is 1.5. We could group our figures, say, into blocks of 5, so long as our highest group includes the value 40.0 and our lowest the value 1.5. So we will have eight groups.

Put the data into this table - some entries are already in:

x	f	
0 - 5	6	
5 - 10	7	5 - 10 means more than 5 and up to 10
10 - 15	10	
* _____	* _____	
* _____	* _____	
* _____	* _____	
* _____	* _____	
* _____	* _____	

Our completed table will look like this:

x	f
0 - 5	6
5 - 10	7
10 - 15	10
15 - 20	8
20 - 25	6
25 - 30	4
30 - 35	3
35 - 40	1
	45 - our check figure

This table we recognise as a frequency distribution table but because it deals with "grouped data" we call this:

GROUPED FREQUENCY DISTRIBUTION TABLE

This is a table where our values are arranged in about 8 groups of equal size - 5 or 10 groups would do, though 10 is better than 5.

If we are looking at say 50 figures all between 0 and 15 a grouping of two would be OK, eg 0 - 2; 2 - 4; and so on.

If we are looking at 50 figures spread between 0 and 700, a grouping of 100 would be a sensible choice 0 - 100; 100 - 200; and so on.

So there are no hard and fast rules to adopt when it comes to choosing group size - we use our judgement guided by some of these ideas we've discussed here.

3. Calculating the mean for grouped data

We have already looked at the mean. Let's remind ourselves of what it is.

SELF CHECK 7
Time guide
2 mins

Put a tick beside the ONE answer which is most correct.

The Mean is best described as:

- The "middle" value in a distribution
- The most frequently occurring value in a distribution
- The total of our values divided by the number of values

The answer to this self check is here in the following words, so you'll need to check your answers carefully. The mean is best described as $\bar{x}$, the total of our values divided by the number of values.

We'll concentrate on the mean because, as we'll see, it has a lot of uses.

But, look back at our completed grouped frequency distribution table for a moment - how do we get our 'fx' column?

Earlier on we multiplied 'x' by 'f' - but we can't do this here because 'x' can be any value between the bottom and top value in each group.

Fortunately our 'mean' comes to the rescue! For example:

in the 0 - 5 group the mean is $2\frac{1}{2}$;

in the 5 - 10 group the mean is $7\frac{1}{2}$; and so on.

SO WHEN WE DEAL WITH GROUPED FREQUENCY OUR
'x' IS THE MEAN GROUP VALUE

SELF CHECK 8
Time guide
5 mins

Here's a grouped frequency distribution table. Some values are entered - complete the table and calculate the mean:

Group	Group mean x	f	fx
0 - 5	$2\frac{1}{2}$	5	$12\frac{1}{2}$
5 - 10	$7\frac{1}{2}$	10	75
10 - 15	* ____	10	* ____
15 - 20	* ____	15	* ____
20 - 25	* ____	20	* ____
25 - 30	* ____	15	* ____
30 - 35	* ____	10	* ____
35 - 40	* ____	10	* ____
40 - 45	* ____	5	* ____
TOTALS		100	* ____

Mean = * ____ divided by * ____ = * ____

Here's our completed table:

Group	Group mean x	f	fx
0 - 5	$2\frac{1}{2}$	5	$12\frac{1}{2}$
5 - 10	$7\frac{1}{2}$	10	75
10 - 15	$12\frac{1}{2}$	10	125
15 - 20	$17\frac{1}{2}$	15	$262\frac{1}{2}$
20 - 25	$22\frac{1}{2}$	20	450
25 - 30	$27\frac{1}{2}$	15	$412\frac{1}{2}$
30 - 35	$32\frac{1}{2}$	10	325
35 - 40	$37\frac{1}{2}$	10	375
40 - 45	$42\frac{1}{2}$	5	$212\frac{1}{2}$
TOTALS		100	2250

As before our mean is total of fx divided by total f.

So,

$$\text{mean} = 2250 \text{ divided by } 100 = 22\frac{1}{2}$$

We can see how much easier it is to calculate the mean this way.

4. Summary

GROUPING DATA - is particularly useful for handling a lot of data

DEFINING GROUPS - groups must be defined carefully eg a group of '5 - 10' means for us more than 5 and up to 10 - of course 10 is included in this group

we aim to divide our data into about 8 groups

GROUP MEANS - our x is the group mean eg in a group '5 - 10' our x is $7\frac{1}{2}$.

PART D THE SPREAD VALUES

1. Introduction

In Part B we looked at three commonly used statistics which can be used to describe data. However, we saw that these values were different. In this section we will look at ways of measuring how different or widely spread our values are. Let's see why we need such a measure when describing data.

2. The Range

ACTIVITY 11
Time guide
2 mins

Look at the following sets of figures of washing machine sales over 10 days.

Number of washing machines
sold by salesman A
in each of 10 days

12
15
20
25
30
30
35
35
38
40

280

Number of washing machines
sold by salesman B
in each of 10 days

4
12
17
18
25
35
30
39
40
60

280

The mean sales level for salesman A is $\frac{280}{10}$ or 28 sales

The mean sales level for salesman B is $\frac{280}{10}$ or 28 sales

The means are the same but the two distributions are different.

Write down your suggestion as to how they differ

*

You may have used another statistic such as the median or mode to describe how they differ. Perhaps you compared the two sets of figures in some other way. However what is important for the work that follows is that salesman A sells a minimum of 12 machines and a maximum of 40. This is a narrower spread of values than those of salesman B who sells a minimum of 4 machines and a maximum of 60.

ACTIVITY 12
Time guide
2 mins

Think of a way to measure how widely spread a set of values is. Write your suggestion in the space below.

*

There are a number of ways of measuring spread. The first way that springs to mind is to take the lowest value from the highest. If we did this we would be measuring the RANGE of our distribution. Let's look at an example.

A small chain of 14 shops in London employs the following number of people:

Number of employees	Number of shops
x	f
1	1
2	2
3	4
4	3
5	2
6	1
7	1
	14

The largest shop employs seven people, and the smallest employs only one. The range is therefore $7 - 1$, or 6, seven being the largest number of employees and one the smallest.

ACTIVITY 13
Time guide
1 min

Can you think of TWO more examples of ranges we could measure which would tell us something else about this chain of shops? I have entered one myself to get you started.

* The range of salaries

*

*

Typical suggestions might have included:

- * The range of sales measured in £s
- * The range of customer complaints in a month
- * The range of sales measured in items sold

Ranges like these would help us describe differences between the shops.

SELF CHECK 9
Time guide
1 min

Let's assume that the number of customer complaints received by each of the shops in a month was:

2; 4; 4; 5; 6; 9; 9; 9; 10; 10; 11; 12; 12; 15

Write down the range of these figures in the space provided below.

The range is _____ complaints.

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. The range is 13 complaints. The highest number of complaints received was 15 and the lowest number was 2. $15 - 2 = 13$.

3. Modifying The Range

Because the range includes the highest and lowest values, it may be measuring the distance between two extreme values which are only likely to appear very occasionally. Let's look at an example.

A firm pays its salesmen a basic wage plus commission of five percent of sales.

The table below shows the commissions earned by each of 100 salesmen in a week.

Commission earned in £s	Number of Salesmen
0 - 10	1
10 - 20	1
20 - 30	20
30 - 40	29
40 - 50	20
50 - 60	25
60 - 70	3
70 - 80	0
80 - 90	0
90 - 100	1
	100

SELF CHECK 10 Fill in the blank below.

Time guide

1 min

The range of commissions earned in this week was £ _____

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. The range of commissions is £100. We do not know if the salesman in the first group earned £0 or £10. We also do not know how much the salesman in the last group earned exactly. But, as our groups run from 0 - 10 and 90 - 100 we must take the lowest and highest levels to make sure that all our salesmen have been included. Our full range is £100. But we can also see from our table that 98 of them got between £20 and £70. 94 of them got between just over £20 and up to £60.

If we were to employ a new salesman, it would be more realistic to tell him that he was likely to get between just over £20 and up to £60 in a week than to say he could get a commission of between £0 and £100.

If we gave a range of figures between which most of our values lie, we could get a more realistic picture. There are a number of suggestions we could make:

* ignore the 4 highest earners, they represent only 4% of the salesmen

* ignore the 2 lowest earners, they represent only 2% of the salesmen.

Let's reproduce our table again and take a closer look at the possibilities.

Commission earned in £s	Number of salesmen
0 - 10	1
10 - 20	1
20 - 30	20
30 - 40	29
40 - 50	20
50 - 60	25
60 - 70	3
70 - 80	0
80 - 90	0
90 - 100	1
TOTAL	100

For convenience, we have dealt with 100 salesmen because this makes it easier to illustrate our points using percentages. Each salesman is then 1% of the total sales force.

If we look at our table, the vast majority of salesmen earn between just over £20 and up to £60 in commission. The percentage of salesmen earning commissions within this range is $20 + 29 + 20 + 25$ which is 94%.

Another way of putting this would be to say that the percentage earning between just over £20 and up to £60 in commission is 94%.

ACTIVITY 14

Time guide
3 mins

Write down ONE advantage for the salesmen of modifying the range in the way we have done here. I've entered one myself to get you started:

- * It is fairer to prospective salesmen to give a range of likely bonus earning

* _____

Other suggestions might be:

- * If the full range is used, a good salesman may feel he is doing badly because he earns say £60 but has been told it's possible to earn up to another £40 more
- * A low performer has a more realistic assessment of the ground needed to make up.

There are other advantages you could have noted for modifying our range to give a more realistic picture.

For instance, the mean often becomes more realistic and representative of the data if it is calculated from the data in the modified range. Of course we need to be careful about the extent to which we modify the range but as long as we include some 90% - 95% of the data, we should be OK.

4. The Mean Deviation

The ranges we've been looking at tell us the limits within which some or all of our values are. What they don't tell us is HOW those values are spread between those limits. Are they mostly grouped towards the lower values? Are they mostly grouped towards the higher values? Are they fairly evenly spread out?

We can find a measure of how data is spread by calculating a special statistic called the mean deviation.

REFERENCE 3
page 48

As Arthur Hall points out in his "Introduction to Statistics" we can also use the mode or the median instead of the mean. However, the mean is the average most usually used for this purpose, so we'll concentrate on the mean here

SELF CHECK 11 Complete the last column of the table below and calculate the mean.
Time guide
2 mins

x	f	fx
1	3	* _____
2	1	* _____
3	4	* _____
4	6	* _____
5	7	* _____
6	4	* _____
	25	* _____

The mean is * _____.

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. The mean is 100 divided by 25 and is 4. The completed table will look like this:

x	f	fx
1	3	3
2	1	2
3	4	12
4	6	24
5	7	35
6	4	24
	25	100

Now that we know the mean, we can find the difference between the mean and each of our 'x' values. We can list the differences under a column heading of 'd'. Let's do this:

SELF CHECK 12 Find the difference between our mean of 4 and each of our x values by filling in the blanks in our table below. I have filled in a couple of lines as a starter:
 Time guide 2 mins

x	f	fx	d
1	2	2	3
2	2	4	2
3	6	18	* _____
4	6	24	* _____
5	6	30	* _____
6	2	12	* _____
	24		

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. Our completed table will look like this:

x	f	fx	d
1	2	2	3
2	2	4	2
3	6	18	1
4	6	24	0
5	6	30	1
6	2	12	2
	24		

Statisticians would head the column 'd' - the DEVIATION. We'll do the same.

Earlier we noted that the frequency column shows us how many times a particular value is recorded.

Now, in order to move on to calculate the mean of these differences from the mean, we must multiply each difference by the frequency with which it occurs. This is the same procedure we've used already to calculate the mean.

SELF CHECK 13 Here's our table again from our last self check. We extend it by adding a fifth column of frequency multiplied by difference; fd. Complete the fd column and put in the total.

Time guide 2 mins

x	f	fx	d	fd
1	2	2	3	6
2	2	4	2	4
3	6	18	1	* _____
4	6	24	0	* _____
5	6	30	1	* _____
6	2	12	2	* _____
TOTALS	24			* _____

So our completed table so far is like this.

x	f	fx	d	fd
1	2	2	3	6
2	2	4	2	4
3	6	18	1	6
4	6	24	0	0
5	6	30	1	6
6	2	12	2	4
TOTALS	24		9	26

The mean difference we are looking for would be found by dividing the total of all our differences which is 26 by the total number of differences which is 9. So 26 divided by 9 is nearly 3.

The mean difference we have just calculated is called the MEAN DEVIATION.

A small number for the mean deviation means that our values are closely grouped.

A large number for the mean deviation means that our values are more widely spread.

The mean deviation is a measure of DISPERSION, or a measure of how widely spread our values are.

SELF CHECK 14 Practice this technique on this data:

Time guide

4 mins

1. Find the mean deviation

x	f	fx	d	fd
1	4	4	* _____	* _____
3	7	21	* _____	* _____
5	10	50	* _____	* _____
7	7	49	* _____	* _____
9	8	72	* _____	* _____
11	4	44	* _____	* _____
TOTALS	40	240		* _____

The mean is * _____

The mean deviation is * _____ divided by * _____ = * _____.

2. The example we used earlier had a mean deviation of nearly 3. Is that distribution more dispersed than the one you have calculated here?

YES/NO

The answer to this self check is here in the following words, so you'll need to check out your answers carefully. Your completed table should look like the one below.

x	f	fx	d	fd
1	4	4	5	20
3	7	21	3	21
5	10	50	1	10
7	7	49	1	7
9	8	72	3	24
11	4	44	5	20
TOTALS	40	240		102

The mean is the total of our values, 240, divided by the total number of values, 40 and is 6. The mean deviation is the total of our differences, 102, divided by the number of differences, 40 and is $2\frac{1}{2}$. The example earlier in our text had a mean deviation of nearly 3. The mean deviation here is $2\frac{1}{2}$. As $2\frac{1}{2}$ is less than 3 then this distribution is less spread out or less DISPERSED.

THE LARGER THE MEAN DEVIATION THE MORE
WIDELY SPREAD OUR DATA IS

5 Summary

In this part we looked at a number of ways of measuring the spread of a frequency distribution. In particular we looked at

- * THE RANGE - which is the difference between the highest and lowest values in a group of values or data
- * MODIFYING THE RANGE - we saw that it can be sensible to use our judgement to modify our range so that we leave out extreme values
- * THE MEAN DEVIATION - is the mean of the differences of our values from the mean
 - the lower the mean deviation the less the dispersion or spread of data

PART E PERFORMANCE CHECKS

1 End Check

Time guide
60 mins

You may now review your understanding of the unit as a whole by completing this group of questions. DON'T GUESS THE ANSWERS - if you're not sure, flick back into the unit and check before responding.

If you are still unsure then put a question mark against your response before proceeding.

TRUE/FALSE

Respond to each statement by writing TRUE or FALSE in the space provided.

- 1 Statistics is concerned with qualitative data. _____
- 2 Tabulation is a technique used to arrange data in a suitable way. _____
- 3 The mean is the most frequently occurring value in a distribution. _____
- 4 The mean is a statistic which can be used to represent a group of widely varying figures. _____
- 5 The mean deviation is a measure of the spread of data. _____
- 6 It is sometimes sensible to modify a range to cut out extreme values. _____

COMPLETION

Respond by completing each sentence with a suitable words or words.

- 7 If we want to calculate the mean wage of a group of people then we would divide the _____ by the number of people.
- 8 The _____ is the difference between the highest and lowest values in a distribution.
- 9 Frequency distribution tables _____ our data.
- 10 We compare two distributions, the first has a mean deviation of 2.1 and the second 0.8, the first distribution is _____ dispersed.
- 11 _____ is another term for spread of data.
- 12 When we have a large number of values we can handle them more easily by drawing up a _____ table.
- 13 When we have a large number of different figures in our data we would be wise to use a _____ distribution table.

MULTIPLE CHOICE

Respond to each question by putting a tick beside the ONE most correct answer.

- 14 Suppose we analyse 10 figures to calculate the mean,
- a) if their total is 210 our mean is 2.1
 - b) if their total is 345 our mean is 3.5
 - c) if their total is 410 our mean is 41
 - d) if their total is 500 our mean is 0.5
- 15 When we are calculating the mean from a frequency distribution table, we multiply our values for x by their frequencies. This is so that:
- a) we can check that we have included all our values in the table
 - b) we include all our values when adding up their total
 - c) we can find the total number of values in our distribution
 - d) we can calculate the total of values more easily
- 16 Which of the following statements best describes the mean
- a) it is the most frequently occurring value
 - b) a figure which can be used to describe a set of values
 - c) it divides our distribution in half
 - d) it is a measure of dispersion
- 17 The mean deviation is calculated from:
- a) the modified range
 - b) the differences between the mean and the data values
 - c) the differences between the mean and the mode
 - d) the extreme values in a distribution

- 18 When making out a grouped frequency distribution table:
- a) make the groups any size you wish
 - b) divide the range into 10 groups
 - c) use between 5 and 10 groups which cover the range of data
 - d) divide the range into group sizes of 5
- 19 Which of the following items can be described as quantitative data:
- a) job satisfaction experienced by members of your work team
 - b) the mood of the members of your section
 - c) scrap rates on your section
 - d) attitudes towards a job enrichment programme
- 20 If we believe our mean could be unreasonably affected by extreme values, we could:
- a) ignore all the extreme values
 - b) modify the range by leaving out some of the extreme values
 - c) use the mean distribution
 - d) calculate the mean from a frequency distribution table
- 21 Here is a set of figures
- 1; 3; 3; 4; 5; 6; 8; 10
- Which statement is true of these figures
- a) the average is 4
 - b) the range is 6
 - c) the mean deviation is 10
 - d) the mean is 5

END CHECK - Checkout

Audio tape
Side 2
Time guide
30 mins

The answers and explanations for these End Checks are on Side 2 of our Audio Tape. Have your responses with you as you play this tape. As you listen you can check out how well you have done and make notes as you proceed.

2 Tutor Check

Time guide
60 mins

A certain machine was tested on each of 20 occasions to find out how many 10 hour shifts it would run without adjustment. The results were as follows:

12.3	14.1	14	13.6	10
2.4	18	20.7	13	20
17.8	8.6	3.2	19.6	12.6
15.5	16.5	18.6	15	9.1

You only need to write ONE or TWO sentences against each question.

- 1) Complete a grouped frequency distribution table of this data.
- 2) Use your table to calculate:
 - a) the mean
 - b) the mean deviation
 - c) the range and suggest a modified range
- 3) A different make of machine is also available.

The mean number of 10 hour shifts is $9\frac{1}{2}$ with a mean deviation of 3.6.

Which is the best machine? The original one or this different make of machine?

Send to
Tutor

Now get your answers to your study tutor either by post or handing it over personally. You will receive your tutor's assessment and comment back through the post as soon as can be managed.

3 Work-based Project

Time guide
60 mins

This is an exercise based on data collected at your place of work.

Make 20 observations, or collect 20 sample figures on any convenient and appropriate area.

Examples might include

- Number of items in stock in twenty periods
 - Number of miles travelled to work by twenty people
 - Output on each of 20 days in your section
 - Number of defects in each of 20 days in your section
- 1) Draw up a frequency distribution table, which may be either grouped or ungrouped depending on the data you collect.
 - 2) Calculate the mean.
 - 3) Calculate the mean deviation and give the range, whether modified or not.
 - 4) Write down TWO conclusions you can draw from these statistics.

Send to
tutor

Now get your answers to your study tutor either by post or by handing it over personally. You will receive your tutor's assessment and comment back through the post as soon as can be managed.

PART F UNIT REVIEW

1. Return to objectives

Having completed your work on this unit let's review how you feel about our four objectives. I'll repeat them here and add brief comment.

Our first objective was

You will be better able to:

* UNDERSTAND THE NATURE AND USES OF STATISTICS AT WORK

We know that statistics may be produced in a number of areas and have seen how they may be used to gain a better picture of what is going on. If this unit has helped you to appreciate the value of numerical information and to better understand what is going on, then you have achieved this objective.

Our second objective was

You will be better able to:

* UNDERSTAND INFORMATION WHICH IS OBTAINED FROM FIGURES AT WORK

Understanding the information that is within the data at work, by reassembling it and calculating statistics from the data, means that you have achieved this objective.

Now to our third objective

You will be better able to:

* USE STATISTICS TO PRESENT AND DESCRIBE NUMERICAL DATA

We have looked at tables, frequency distribution tables in particular and seen how to describe sets of figures using the mean and measures of spread. Having done the work in the unit, assembling and calculating various statistics, you have met this objective.

Finally, our fourth objective

You will be better able to:

* FORM JUDGEMENTS BY 'READING' VARIOUS KINDS OF DATA PRESENTATIONS AND STATISTICS

This objective is an extension of the last one. Now that you can better present and describe numerical data yourself and are able to use this knowledge to draw certain conclusions, then you have achieved this objective.

I hope that the work you have done was enjoyable and useful enough to encourage you to continue in your studies.

2 Extensions

I recommend that you take up as many of these extensions as you can. They will further increase your understanding and interest. You will also be able to check out a number of the ideas that I mention in the unit. The extra time and effort will prove very worthwhile.

EXTENSION 1 Book: An Introduction to Statistics
 Author: A Hall
 Publisher: The Macmillan Press

EXTENSION 2 Book: Using Statistics in Business 1
 Author: Peter Clark
 Publisher: Pan - Breakthrough series

Chapter 7 gives a fuller discussion of approximations than we were able to give in this unit.

EXTENSION 3 Book: Choice Series Unit 32: Using Graphs
 Author: Linda Carr
 Publisher: Northern Regional Management Centre

EXTENSION 4 Book: Supervisory Studies
 Author: P W Betts
 Publisher: Macdonald and Evans

Chapter 24 provides a good account of some of the graphical and diagrammatic methods of presentation not covered in this unit.

These extension opportunities can all be taken up via your Support Centre.

They will either have the materials or will arrange that you have access to the materials. However it may be more convenient to check out the materials with your Personnel or Training people at work - they may well give you access. There are other good reasons for approaching your own people, for example they will become aware of your interest and you can involve them in your development.

3 References

- REFERENCE 1 Book: Using Statistics in Business 1
 Author: Peter Clark
 Publisher: Pan - Breakthrough series
- REFERENCE 2 Book: Supervisory Studies
 Author: P W Betts
 Publisher: Macdonald and Evans
- REFERENCE 3 Book: An Introduction to Statistics
 Author: A Hall
 Publisher: The Macmillan Press

These references can all be taken up via your Support Centre. They will either have the materials or will arrange that you have access to the materials. However check with your Personnel or Training people at work - they may well give you access. There are other good reasons for approaching your own people, for example they will become more aware of your interest and you can involve them in your development.

If you do wish to make use of these references then don't think that you must read them right through. Just dip into them for a slightly wider treatment of the topic raised by the unit.

I should warn that the language and style of the writing is often difficult to cope with. So don't struggle with it too much. Do what I do when in difficulty, put the book back on the shelf and find a better one!

THE CHOICE SERIES

PRINCIPLES AND PRACTICE OF SUPERVISION

- 11 Needs and Rewards
- 12 Enriching Work
- 13 Workteams
- 14 Team Leading
- 15 Leading Change
- 16 Organisation Systems
- 17 Supervising in the System
- 18 Supervising with Authority
- 19 Workplace Control
- 20 Taking Decisions

TECHNICAL ASPECTS OF SUPERVISION

- 30 Looking at Figures
- 32 Using Graphs
- 33 Method Study
- 34 Easy Statistics
- 35 Quality Control
- 37 Controlling Output
- 38 Value Analysis
- 39 Quality Circles
- 40 Computers
- 41 Stores Control
- 42 Managing Time
- 43 Descriptive Statistics
- 45 Supervisors & Marketing

COMMUNICATION

- 60 Communicating
- 61 Speaking Skills
- 62 Writing Skills
- 63 Communication Systems
- 64 Orders and Instructions

ECONOMIC AND FINANCIAL ASPECTS

- 70 Accounting for Money
- 71 Control via Budgets
- 72 Cost Reduction
- 73 Wage Payment Systems
- 74 The National Economy

INDUSTRIAL RELATIONS

- 80 Training Plans
- 81 Training Sessions
- 82 Discipline and the Law
- 84 Health and Safety
- 85 Industrial Relations in Action
- 86 Equality at Work
- 87 Hiring People
- 88 Supervising and the Law